A UT/ORNL PARTNERSHIP
NATIONAL INSTITUTE FOR COMPUTATIONAL SCIENCES
NICS

Kraken Overview

Daniel Lucio
User Support

March 9th 2011, ORNL, TN

National Institute for Computational Sciences

A UT/ORNL PARTNERSHIP
NATIONAL INSTITUTE FOR COMPUTATIONAL SCIENCES
NICS

- NICS is a collaboration between UT and ORNL
- Awarded the NSF Track 2B ($65M)
- Staffed with 25 FTEs

Kraken's Timeline

NSF grant awarded in late '07

	XT3	XT4	Initial XT5	Pre-XT5	Final XT5
	April '08	July '08	Feb '09	Oct '09	Feb '11
Compute Cores	7,352	18,048	66,048	99,072	112,896
Compute Memory	7.4TB	17.6TB	100TB	129TB	147TB
# Cabinets	40	48	88	88	100
Peak FLOPS	38.6TF	166.5TF	608TF	1.03PF	1.17PF
Top500 Ranking	#57	#15	#6	#3	?

8th Most Powerful SuperComputer

TOP500 List - November 2010 (1-100)

R_{max} and R_{peak} values are in TFlops. For more details about other fields, check the TOP500 description.

Power data in KW for entire system

next

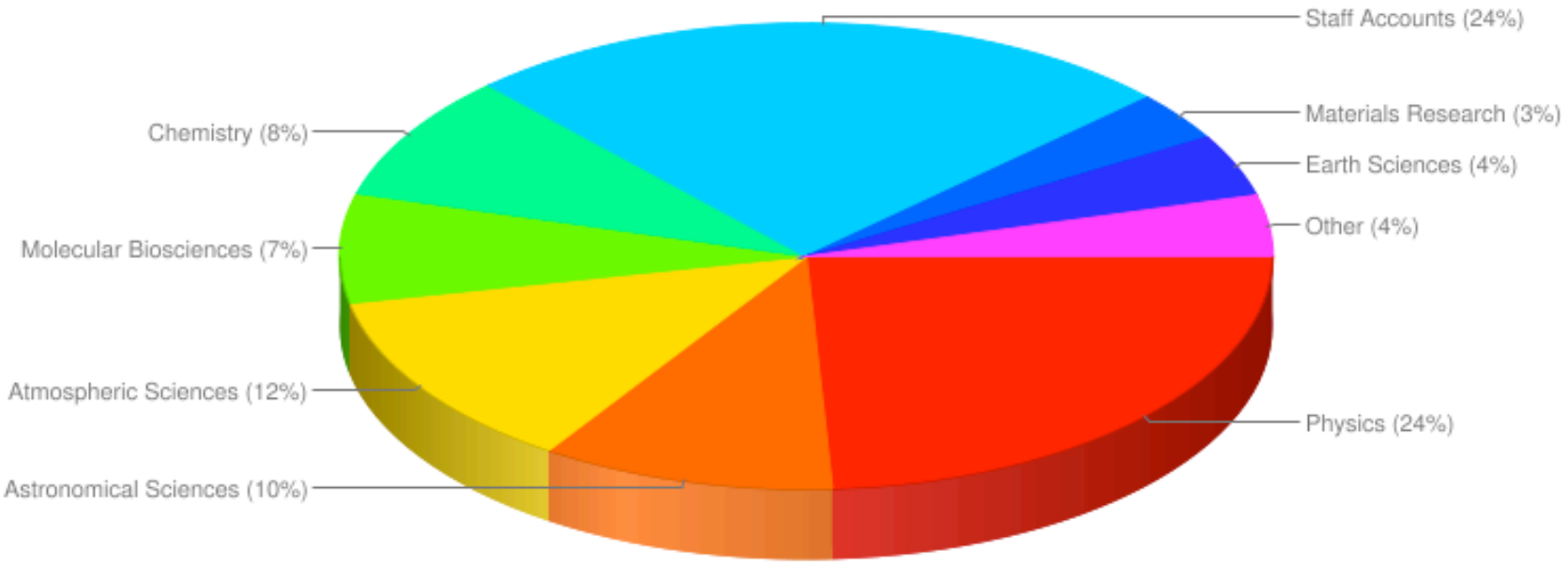
Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	National Supercomputing Center in Tianjin China	Tianhe-1A - NUDT TH MPP, X5670 2.93Ghz 6C, NVIDIA GPU, FT-1000 8C / 2010 NUDT	186368	2566.00	4701.00	4040.00
2	DOE/SC/Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron 6-core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
3	National Supercomputing Centre in Shenzhen (NSCS) China	Nebulae - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU / 2010 Dawning	120640	1271.00	2984.30	2580.00
4	GSIC Center, Tokyo Institute of Technology Japan	TSUBAME 2.0 - HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows / 2010 NEC/HP	73278	1192.00	2287.63	1398.61
5	DOE/SC/LBNL/NERSC United States	Hopper - Cray XE6 12-core 2.1 GHz / 2010 Cray Inc.	153408	1054.00	1288.63	2910.00
6	Commissariat a l'Energie Atomique (CEA) France	Tera-100 - Bull bullx super-node S6010/S6030 / 2010 Bull SA	138368	1050.00	1254.55	4590.00
7	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
8	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron 6-core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	3090.00
9	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
10	DOE/NNSA/LANL/SNL United States	Cielo - Cray XE6 8-core 2.4 GHz / 2010 Cray Inc.	107152	816.60	1028.66	2950.00

Largest Teragrid resource

High Performance Systems

Name	Institution	System	Peak TFlops	Memory TBytes	Status	Load	Running Jobs	Queued Jobs	Other Jobs
Kraken	NICS	Cray XT5	1030.00	129.00	Down		0	748	255
Ranger	TACC	Sun Constellation	579.40	123.00	Up		254	885	246
Lonestar	TACC	Dell Linux Cluster	302.00	45.00	Up		209	1	10
Athena	NICS	Cray XT4	166.00	17.60	Up		80	20	3
Abe	NCSA	Dell Intel 64 Linux Cluster	89.47	9.38	Up		263	276	160
Steele	Purdue	Dell Intel 64 Linux Cluster	60.00	12.40	Up		695	85	174
Queen Bee	LONI	Dell Intel 64 Linux Cluster	50.70	5.31	Up		45	7	0
Lincoln	NCSA	Dell/Intel PowerEdge 1950	47.50	3.00	Up		34	8	0
Big Red	IU	IBM e1350	30.60	6.00	Up		165	58	1
Frost	NCAR	IBM BlueGene/L	22.90	2.00	Up		20	4	0
Ember	NCSA	SGI Altix UV	16.00	8.00	Down		0	55	5
Pople	PSC	SGI Altix 4700	5.00	1.54	Up*		11	16	22
Dash	SDSC	Intel Nehalem	4.90	3.00	Up				
NSTG	ORNL	IBM IA-32 Cluster	0.34	0.07	Up		0	0	1
		Total:	**2504.81**	**385.55**			**1776**	**2163**	**877**

Actual usage by discipline (Feb ‘11)



Total Projects: 783

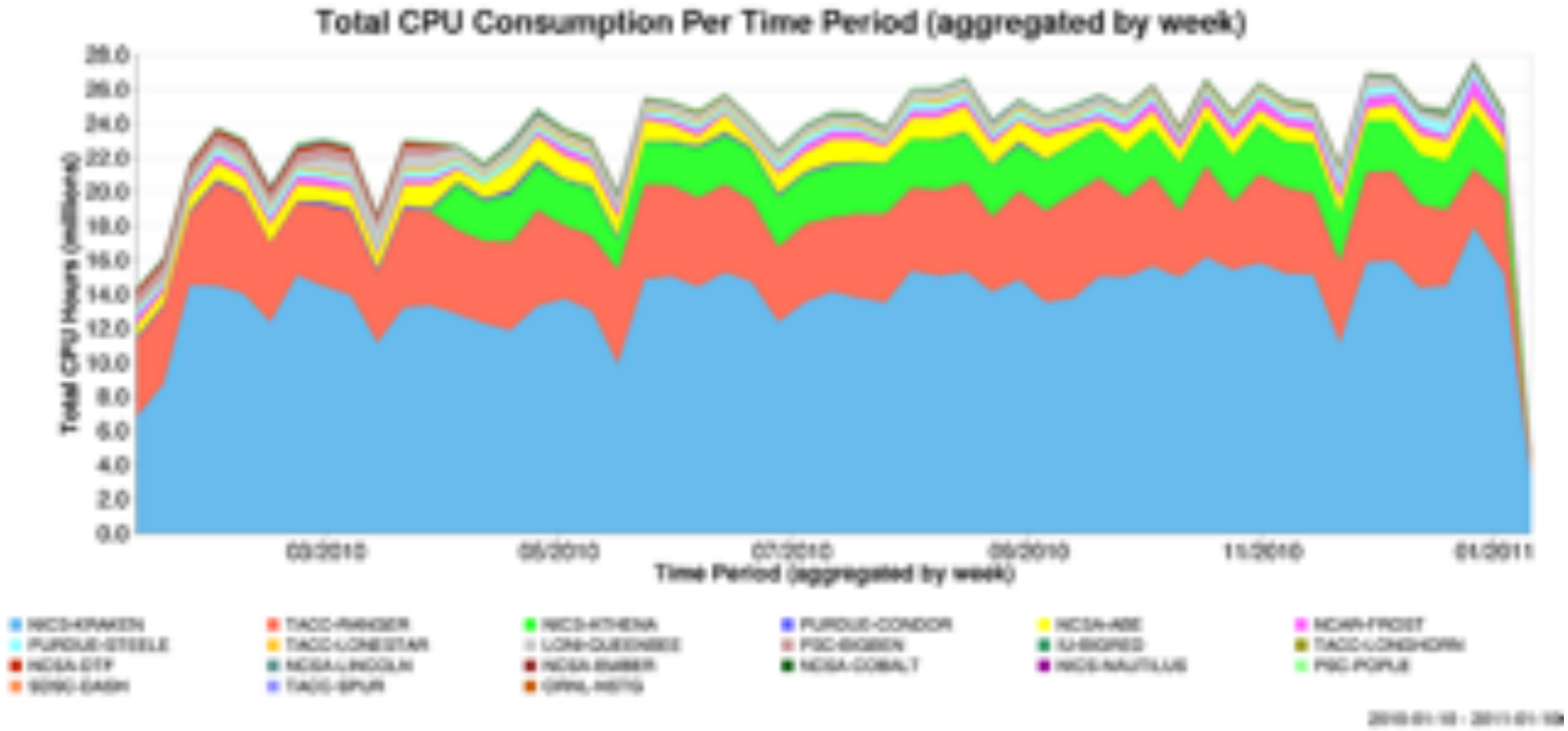
TG:~577 + UT:~46 + DD:160

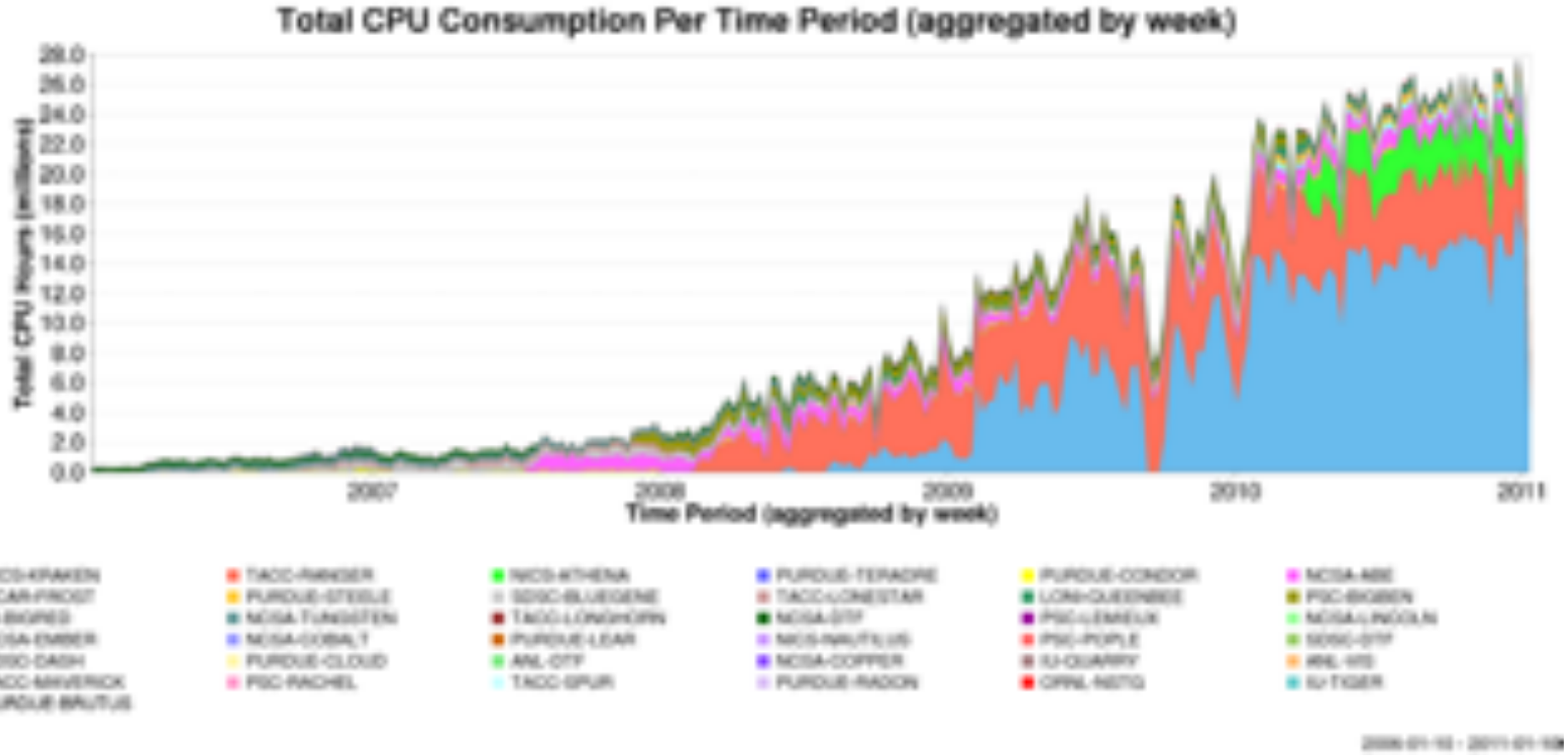
Total Users: ~3,365

Active Users: ~800

Allocated 650M S.U. hours in ‘10

80% for TG – 20% for UT


Total CPU Consumption Per Time Period (aggregated by week)
Total CPU Hours (millions)
28.0
26.0
24.0
22.0
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
03/2010
05/2010
07/2010
09/2010
11/2010
01/2011
Time Period (aggregated by week)
NICS-KRAKEN
TACC-RANGER
NICS-ATHENA
PURDUE-CONDOR
NCSA-ABE
NCAR-FROST
PURDUE-STEELE
TACC-LONESTAR
LONI-QUEENBEE
PSC-BIGBEN
IU-BIGRED
TACC-LONGHORN
NCSA-ETF
NCSA-LINCOLN
NCSA-SILVER
NCSA-COBALT
NICS-NAUTILUS
PSC-POPLE
SDSC-DASH
TACC-SPUR
GRNL-NSTG


Total CPU Consumption Per Time Period (aggregated by week)
Total CPU Hours (millions)
28.0
26.0
24.0
22.0
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
2007
2008
2009
2010
2011
Time Period (aggregated by week)
NICS-KRAKEN
TACC-RANGER
NICS-ATHENA
PURDUE-TERAGRE
PURDUE-CONDOR
NCSA-ABE
NCAR-FROST
PURDUE-STEELE
SDSC-BLUEGENE
TACC-LONESTAR
LONI-QUEENBEE
PSC-BIGBEN
IU-BIGRED
NCSA-TUNGSTEN
TACC-LONGHORN
NCSA-DTF
PSC-LEMIEUX
NCSA-LINCOLN
NCSA-EMBER
NCSA-COBALT
PURDUE-LEAR
NICS-NAUTILUS
PSC-POPLE
SDSC-DTF
SDSC-DASH
PURDUE-CLOUD
ANL-DTF
NCSA-COPPER
IU-QUARRY
ANL-VIS
TACC-MAVERICK
PSC-RACHEL
TACC-GPUR
PURDUE-RADON
ORNL-NSTG
IU-TIGER
PURDUE-BRUTUS


Kraken Job Mix Jan 2010 to Jan 2011
25,000,000
20,000,000
15,000,000
10,000,000
5,000,000
0
Hours
Jan-11
Dec-10
Nov-10
Oct-10
Sep-10
Aug-10
Jul-10
Jun-10
May-10
Apr-10
Mar-10
Feb-10
Jan-10
Month
63
127
255
511
1024
2047
4096
8192
16384
Node Groupings

Kraken System Configuration

- Cray XT5 running CLE 2.2UP02 (soon 2.2UP03)
- 100 cabinets in 4 rows
- 9,408 compute nodes (112,896 cores) & 96 service nodes
- 147TB of compute memory
- Two file systems available
  - NFS mounted home areas, 2TB
  - Lustre Scratch space, with 2.4PB of usable space
- 25x16x24 3D torus topology interconnect using SeaStar2 chips

Compute node configuration

- Two 2.6 Ghz Six-Core AMD (Istanbul) Processors
- Dual socket – 12 cores per node
- 16GB RAM per node
- Diskless nodes
- The ONLY accessible file system is Lustre scratch
- Runs a streamlined version of Linux-like OS called CLE
- Users cannot login to the compute nodes
- You need `qsub` & `aprun` to launch jobs in these nodes
- TORQUE/MOAB & ALPS control these resources

Service node configuration

- One 2.6 Ghz Dual-Core AMD Processors
- One socket – 2 cores per node
- 8GB RAM per node
- Diskless nodes
- Both NFS home areas & Lustre scratch accessible
- Runs a complete Linux-like OS called SLES10SP1
- There are 16 login nodes
- 11 OTP only + 4 GSISSH only + 1 Experimental
- 4 GridFTP only with 10GigE internet connection
- 16 Aprun nodes & 48 I/O nodes

Important Policies

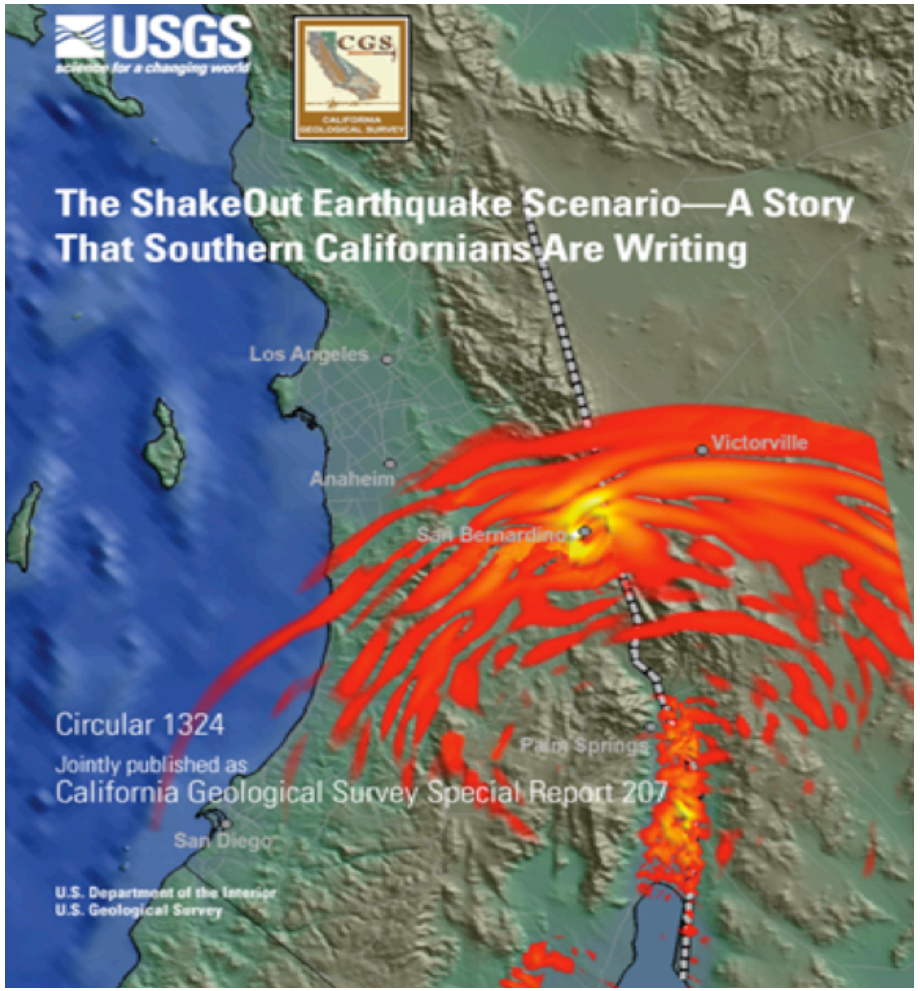
- No production jobs should be run at the login (service) nodes
- Jobs using an account with a negative balance will run only as backfill jobs
- Large core count (i.e. capability) jobs have more priority
- Dedicated mode of the whole system is possible on Wednesdays
- Refunds can be provided for up to 6hrs
- When Lustre gets 70% full we contact users to ask them to delete files. When 80% full, we will start deleting oldest files as an emergency procedure

Important Changes since last time

- Hardware upgrade from 99,072 to 112,896 cores.
- Intel compiler is now available (Cray's soon)
- TORQUE upgrade. After this upgrade, jobs will die at the beginning if output files cannot be created.
- Refunds can be provided for at most 6hrs. Checkpointing is important.
- 'longsmall' queue has been deactivated. Max walltime for jobs less than 49K cores is 24hrs.

Simulating “The Big One”

- Performed the largest earthquake simulation ever on the San Andreas Fault on Kraken
- Simulated in a 32 billion grid point subset of the SCEC Community Velocity Model (CVM) V4
- Used 96,000 processor cores



Slide info courtesy of Phil Maechling, Southern California Earthquake Center

NICS

Cosmology Simulations of the Lyman Alpha Forest

- Performed the largest hydrodynamic cosmology simulation ever done on Kraken
- Used ENZO (Hybrid MPI/ OpenMP code) for current model of $4{,}096^3$ = 64 billion dark matter particles
- *"The most productive platform in NSF portfolio for ENZO simulations, bar none."*

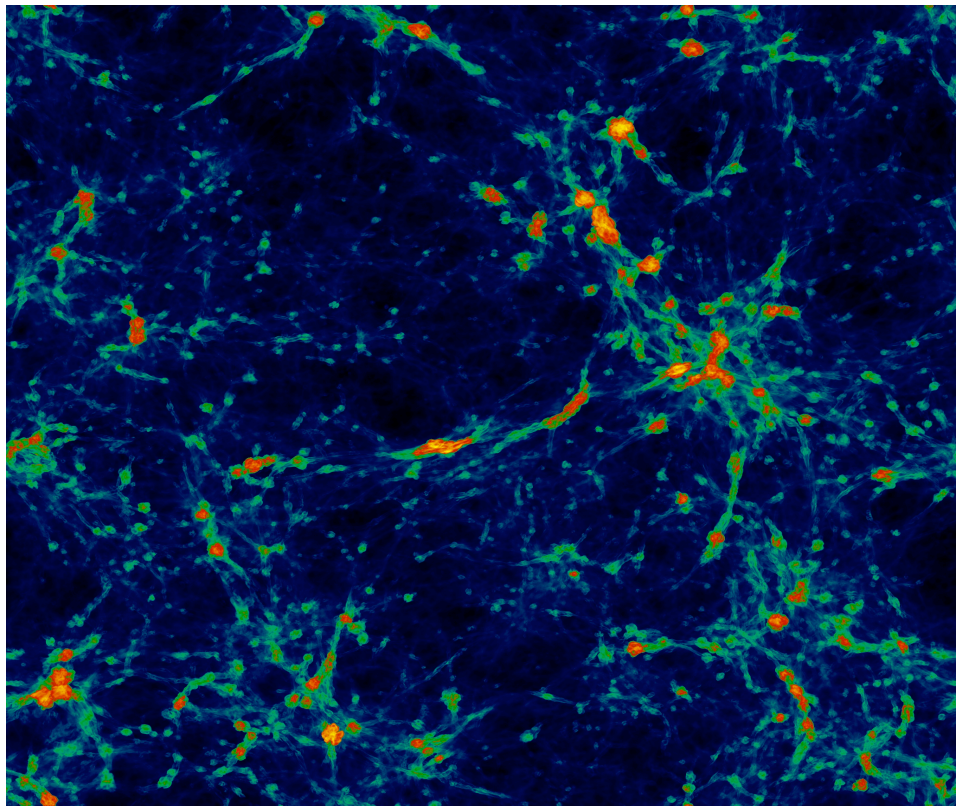


Image of the Lyman Alpha Forest showing the Baryon Acoustic Oscillation (BAO), which arises from sound waves becoming "frozen" when the matter and radiation decouple in the Big Bang.

Slide info courtesy of Robert Harkness, University of California, San Diego

Other NICS HPC resources

For more information on other NICS HPC resources, please visit

http://www.nics.tennessee.edu/computing-resources
http://rdav.nics.tennessee.edu/resources
http://keeneland.gatech.edu/